

AD 681065

ON THE INITIAL VALUE PROBLEM FOR  
ROTATING STRATIFIED FLOW

By LOUIS N. HOWARD AND WILLIAM L. SIEGMANN

1. Introduction. In this paper we shall investigate some general properties of motions of a compressible fluid under the influence of a gravitational field. Our main restrictive hypotheses are first that we are concerned with motions (or time scales) for which the dissipative effects of viscosity and heat condition are not important, and second that the motions are small deviations from a basic state of rigidly rotating hydrostatic equilibrium. We also assume that the gravitational field is externally given, and is not itself affected by the motion. The mathematical model based on these hypotheses is the linear theory of rotating stratified non-dissipative flow. Perhaps the most basic mathematical problem in this context is the initial-boundary value problem: the values of the flow variables are given at some initial time in a region  $R$ , and their subsequent temporal evolution is sought, subject to appropriate boundary conditions on the boundary  $B$  of  $R$ . We shall be concerned with this problem in the case of the boundary condition corresponding to a rigid envelope  $B$ , i.e. we take the normal component of the velocity vector to be zero on  $B$ . As usual with initial-boundary value problems for linear systems of partial differential equations with time independent coefficients, it is helpful to think of the problem in terms of the "method of normal modes," a normal mode being

JAN 22 1969

a solution of the differential equations which satisfies the (homogeneous) boundary conditions and varies harmonically with time. If the initial data can be represented as a superposition of the initial values of the various normal modes, then the solution to the initial-boundary value problem is simply the same superposition of normal mode components, each evolving in time at its own frequency. From this point of view there are two parts to the problem: first the determination of the frequencies and the spatial dependence of the various normal modes, and second the representation of the initial data in terms of their initial values. In many problems, for instance in heat conduction, the first part is by far the more difficult, the representation of initial data being readily found by exploiting some sort of spatial orthogonality of normal modes of different frequencies. To a certain extent this is true of the present problem - it is only in exceptional cases that it is possible to describe the oscillatory normal modes explicitly (indeed even their existence is not an easy question), though they do have a kind of spatial orthogonality. However in the case of steady flows, normal modes of frequency zero, it turns out to be possible to describe them quite generally. But in the present case, as in the analogous problem for homogeneous rotating flow, this class of steady flows is often quite extensive - the frequency zero is highly degenerate, and this means that the second part of the problem - the determination of exactly which steady component is excited by the initial data - is by no means trivial. Our main results deal with this aspect of the problem. After formulating the problem and describing some of its general features,

we shall describe all the steady flows. Then we shall establish some basic conservation properties possessed by all flows in this linear theory, and show how they can be used to characterize exactly, from the initial data, the steady component of the solution to the initial-boundary value problem. We conclude with some examples illustrating our general results in special cases, and compare the present problem with its analogue for rotating homogeneous flow.

2. Formulation. The basic equations of our mathematical model are those of a compressible fluid without viscosity or heat conduction, written relative to a system rotating about the z-axis with angular velocity  $\Omega$ :

$$\underline{u}_t + 2\Omega \underline{k} \times \underline{u} + \underline{u} \cdot \nabla \underline{u} + \frac{1}{\rho} \nabla p + \nabla \varphi = 0 \quad (1)$$

$$\rho_t + \nabla \cdot (\rho \underline{u}) = 0 \quad (2)$$

$$s_t + \underline{u} \cdot \nabla s = 0 \quad (3)$$

$$\rho = f(p, s) \quad (4)$$

Here  $\underline{u}$ ,  $\rho$ ,  $p$  and  $s$  are velocity, density, pressure and specific entropy and  $\varphi = -\Omega^2(x^2 + y^2)/2 + \varphi_1$  is the "geopotential,"  $\varphi_1$  being the ordinary gravitational potential; in the simplest case we would have  $\varphi_1 = gz$ , but we do not require any specific form for  $\varphi_1$ , nor need we specialize the equation of state (4), though we shall assume

(as is appropriate for all ordinary fluids) that  $\partial f / \partial p \geq 0$ . Let  $L$  be a length characterizing the size of the region  $R$  containing the fluid, and let  $\tilde{\rho}$  and  $\tilde{s}$  be characteristic density and entropy scales. Then introduce dimensionless variables, indicated by asterisks, as follows:  $\underline{r} = L\underline{r}_*$ ,  $t = \Omega^{-1}t_*$ ,  $\underline{u} = \Omega L\underline{u}_*$ ,  $\rho = \tilde{\rho}\rho_*$ ,  $s = \tilde{s}s_*$ ,  $p = L^2\Omega^2\tilde{\rho}p_*$ ,  $\varphi = \Omega^2L^2\tilde{\rho}\varphi_*$ . The dimensionless equations are then the same as (1) - (4) except that  $\Omega$  is replaced by 1, asterisks are affixed to all the variables, and for the equation of state we write  $f_*(s_*, p_*) = \tilde{\rho}^{-1}f(\tilde{s}s_*, L^2\Omega^2\tilde{\rho}p_*)$ . Let the basic state about which we shall linearize have  $\underline{u}_* = 0$ ,  $\rho_* = \rho_0$ ,  $p_* = p_0$ ,  $s_* = s_0$ ; the equations then show that  $\rho_0$ ,  $p_0$  and  $s_0$  are all functions of  $\varphi_*$ , and in fact  $\rho_0 = -p'_0(\varphi_*)$  and  $\rho_0 = f_*(s_0, p_0)$ . We may thus consider the basic state of stratification to be specified by, say,  $p_0(\varphi_*)$ , and  $\rho_0$  and  $s_0$  are then determined. It should be noted that while we shall usually write expressions like  $\rho_0 = \rho_0(\varphi_*)$ , this should not be interpreted quite literally, for what the equations for the basic state really say is that  $\rho_0$  does not change as one moves along a surface of constant  $\varphi_*$  in the flow region. But the constant  $\varphi_*$  surface may be cut into several disconnected components by the flow region, and though  $\rho_0$  must be a constant within each component, the different components may have different values of this constant. The density distributions down inside each of two potholes in the bottom of a stratified lake need not agree. The linearized equations are now obtained by setting  $\underline{u}_* = \varepsilon\underline{u}_1$ ,  $\rho_* = \rho_0 + \varepsilon\rho_1$ ,  $p_* = p_0 + \varepsilon p_1$ ,  $s_* = s_0 + \varepsilon s_1$  in (1) - (4) and retaining first order terms in  $\varepsilon$ ;  $\varepsilon$  may be regarded as a sort of Rossby number characterizing the

magnitude of the initial perturbation. Suppressing from now on the subscripts 1 and \*, and writing  $f_s$  and  $f_p$  for  $\partial f_*/\partial s_*$  and  $\partial f_*/\partial p_*$  evaluated at the basic state, these linearized equations are:

$$\underline{u}_t + 2\underline{k} \times \underline{u} + \rho_0^{-1} \nabla p + \rho_0^{-1} \rho \nabla \varphi = 0 \quad (5)$$

$$\rho_t + \nabla \cdot (\rho_0 \underline{u}) = 0 \quad (6)$$

$$s_t + \underline{u} \cdot \nabla s_0 = 0 \quad (7)$$

$$\rho = f_s s + f_p p. \quad (8)$$

These equations are to be considered in a fixed region  $R$ , with the boundary condition that  $\underline{u} \cdot \underline{n} = 0$  on the boundary  $B$  of  $R$ ,  $\underline{n}$  being the unit outward normal to  $B$ . For the initial value problem we also prescribe  $\underline{u}$ ,  $\rho$  and  $s$  at  $t = 0$ .

The equations (5) - (8) possess an invariant (i.e. temporally conserved) "energy integral." It can be obtained as follows:

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{1}{2} \rho_0^{-1} f_p p^2 \right) &= \rho_0^{-1} p (\rho_t - f_s s_t) \\ &= \rho_0^{-1} p (-\rho_0 \nabla \cdot \underline{u} - \rho_0' \underline{u} \cdot \nabla \varphi + s_0' f_s \underline{u} \cdot \nabla \varphi) \\ &= -p \nabla \cdot \underline{u} - \rho_0^{-1} p f_p p_0' \underline{u} \cdot \nabla \varphi = -p \nabla \cdot \underline{u} + p f_p \underline{u} \cdot \nabla \varphi, \end{aligned}$$

since  $\rho_0' = f_s s_0' + f_p p_0'$  from the original equation of state. But we also have

$$\frac{\partial}{\partial t} \left( \frac{1}{2} (s'_0)^{-1} f_s s^2 \right) = -(s'_0)^{-1} f_s s s'_0 \underline{u} \cdot \nabla \varphi = -s f_s \underline{u} \cdot \nabla \varphi,$$

and so

$$\begin{aligned} \frac{\partial}{\partial t} \left( \frac{1}{2} \rho_0^{-1} f_p p^2 - \frac{1}{2} (s'_0)^{-1} f_s s^2 \right) &= -\underline{p} \nabla \cdot \underline{u} + \rho \underline{u} \cdot \nabla \varphi \\ &= -\nabla \cdot (\underline{p} \underline{u}) + \underline{u} \cdot (\nabla \underline{p} + \rho \nabla \varphi) \\ &= -\nabla \cdot (\underline{p} \underline{u}) - \frac{1}{2} \frac{\partial}{\partial t} (\rho_0 |\underline{u}|^2) \end{aligned}$$

using (5). Thus

$$\frac{\partial}{\partial t} \frac{1}{2} \left\{ \rho_0 |\underline{u}|^2 + \rho_0^{-1} f_p p^2 - (s'_0)^{-1} f_s s^2 \right\} + \nabla \cdot (\underline{p} \underline{u}) = 0. \quad (9)$$

Integrating this over  $R$  and using the boundary condition we find the energy conservation formula:

$$\frac{\partial \mathcal{E}}{\partial t} \equiv \frac{\partial}{\partial t} \frac{1}{2} \int_R \left\{ \rho_0 |\underline{u}|^2 + \rho_0^{-1} f_p p^2 - (s'_0)^{-1} f_s s^2 \right\} dV = 0. \quad (10)$$

We assume that  $f_p \geq 0$  (density increases with pressure if no heat is added), and we shall also suppose that the basic state is such that  $(s'_0)^{-1} f_s < 0$ , i.e. it is "statically stable." Usually this would mean  $s'_0 > 0$  (specific entropy increases with altitude), since most ordinary fluids expand on being heated at constant pressure, i.e.  $f_s < 0$ . With these assumptions we see that the quadratic form  $\mathcal{E}$  is positive definite and thus the perturbation quantities will remain bounded (in  $L_2$  norm anyway) if they are so initially. It

is of course to achieve this, i.e. to rule out the possibility of convective instability, that we assume static stability of the basic state.

If we introduce the inner product corresponding to the quadratic form  $\mathcal{E}$  by

$$\langle \underline{u}_1, p_1, s_1 | \underline{u}_2, p_2, s_2 \rangle = \int_R [\rho_0 \underline{u}_1^* \cdot \underline{u}_2 + \rho_0^{-1} f_p p_1^* p_2 - (s_0')^{-1} f_s s_1^* s_2] dV \quad (11)$$

(the asterisk denotes the complex conjugate) one readily verifies that the inner product of two (possibly complex) solutions of (5) - (8) is also invariant in time. In particular this implies the reality of the frequency of any normal mode solution of (5) - (8), and the orthogonality, in the sense of the inner product (11), of any two normal modes of different frequencies.

We note also that the energy conservation formula (10) implies the uniqueness of the solution to the initial value problem, for the difference of two solutions with the same initial data is also a solution of the equations and the boundary condition, and has zero energy initially. Being conserved, its energy is also zero thereafter, and since it is positive definite this implies that the difference flow is always zero, i.e. the original solution is unique. A proof of the existence of a solution to the initial value problem appears to be a more difficult question - we shall merely assume it here.

3. Steady flows. The time-independent solutions of equations (5) - (8) can be easily described explicitly. First of all, (7) shows that the velocity field of any such solution must be horizontal, i.e. directed along surfaces of constant  $\varphi$ ,  $\underline{u} \cdot \nabla\varphi = 0$ . (We shall always use the word "horizontal" in this sense - vectors perpendicular to  $\underline{k}$  are thus not usually horizontal.) If this is true, then the cross product of  $\nabla\varphi$  with the time independent form of (5) shows that

$$\underline{u} = (2\varphi_z \rho_0)^{-1} \nabla\varphi \times \nabla p \quad (12)$$

and putting this back in (5) we find that  $\rho$  must be given by

$$\rho = - \frac{p_z}{\varphi_z} . \quad (13)$$

These two necessary conditions together with the equation of state (8) give us a candidate for a steady solution of (5) - (8) for any function  $p$ , and in fact one easily checks that formulas (12) and (13) are sufficient for (5) and (7) to be satisfied. However (6) will be satisfied only if also  $\nabla p \cdot \nabla\varphi \times \nabla\varphi_z = 0$ ; this is no restriction on  $p$  in case  $\nabla\varphi \times \nabla\varphi_z \equiv 0$ , but otherwise it shows that  $p$  must be constant along the integral curves of the vector field  $\nabla\varphi \times \nabla\varphi_z$ , i.e. along the intersections of the surfaces of constant  $\varphi$  and those of constant  $\varphi_z$ . Parts of the flow region  $R$  throughout which  $\nabla\varphi \times \nabla\varphi_z = 0$ , i.e. in which  $\varphi$  and  $\varphi_z$  are functionally dependent, we shall call geostrophically free regions.

Elsewhere, we imagine the integral curves of  $\nabla\varphi \times \nabla\varphi_z$  to be constructed, and call them geostrophic curves. Through each point in parts of  $R$  which are not geostrophically free (if any) there passes a unique geostrophic curve; some or all of these curves may cross the boundary of the region  $R$ . Those parts of  $R$  which are covered by geostrophic curves which cross the boundary  $B$  we call geostrophically blocked regions. Other parts of  $R$  which are covered by geostrophic curves which do not cross the boundary we call geostrophically guided regions. We shall assume that the geostrophic curves in guided regions are closed, neglecting the possibility that they might wind around infinitely inside the region. It is certainly difficult to imagine a potential function which is at all reasonable for which the geostrophic curves stay in the flow domain but are not closed. With this terminology we may say then that for a steady flow it is necessary that  $p$  be constant along geostrophic curves in guided or blocked regions, i.e. (essentially)  $p = p(\varphi, \varphi_z)$  there. There is an additional minor restriction on  $p$  - it must be such that (12) and (13) do not make  $\underline{u}$  or  $\rho$  singular at places where  $\varphi_z = 0$ , if there are any such. But except for this, (12), (13) and the equation of state (to give  $s$ ) provide the general solution of the steady forms of (5) - (8), with  $p$  arbitrary in free regions, or an arbitrary function constant on geostrophic curves otherwise. However we are actually interested only in solutions which also satisfy the boundary condition  $\underline{u} \cdot \underline{n} = 0$  on  $B$ ; (12) shows that this is always true on horizontal parts of  $B$ , if any, but that it is necessary (and sufficient) that  $p$  be constant along horizontal curves on the non-horizontal part of  $B$ .

Such a steady solution of (5) - (8) which satisfies  $\underline{u} \cdot \underline{n} = 0$  on  $B$  we call a geostrophic flow. In non-free regions we note that

$$\underline{u} = (2\varphi_z \rho_0)^{-1} \frac{\partial p(\varphi, \varphi_z)}{\partial \varphi_z} \nabla \varphi \times \nabla \varphi_z,$$

so that the velocity vector of a geostrophic flow is directed along the geostrophic curves. Furthermore the magnitude of  $\underline{u}$  along a given geostrophic curve (in the non-free case) is a constant multiple of  $|\nabla \varphi \times \nabla \varphi_z|$ , and thus if a geostrophic curve crosses the boundary the velocity vector on it must be zero. In this case the momentum equation (or (12)) shows that  $p$  must not only be constant along geostrophic curves, but constant on horizontal surfaces, i.e. (essentially)  $p = p(\varphi)$  in blocked regions; and (13) and (8) then show that  $\rho$  and  $s$  are likewise constant on horizontal surfaces. Thus a geostrophic "flow" in a blocked region is at most only a perturbation of the basic state into another state of hydrostatic equilibrium.

Summarizing these results, the geostrophic flows are given by (12), (13) and (8) with  $p$  a function which is: (a) arbitrary in free regions, except that it must be constant along horizontal curves on the non-horizontal part of  $B$ , (b) an arbitrary function constant along geostrophic curves in guided regions, or (c) an arbitrary function constant on horizontal surfaces in blocked regions.

4. Conservation theorems. The temporal conservation of the energy integral is a result of a very general sort; something like it is true of any non-dissipative system. We come now to some results which are in a sense much stronger, and relate much more specifically to the properties of linearized rotating stratified flow. The first of these has to do with quantities on the boundary  $B$  of the flow region. Some of  $B$  may consist of pieces of constant  $\varphi$  surfaces - we call this the horizontal part  $B_h$  of  $B$ ; on it  $\underline{n} \times \nabla\varphi = 0$ . Parts of the boundary on which  $\underline{n} \times \nabla\varphi \neq 0$  make up the non-horizontal part  $B_n$ . (Edges, where  $B_h$  and  $B_n$  meet, we may count as part of  $B_h$ .) Since  $\underline{n}$  and  $\nabla\varphi$  have the same direction on the "flat spots"  $B_h$ , the boundary condition and the entropy equation (7) show that  $s_t = 0$  there. Thus entropy is conserved pointwise on flat spots, for any solution satisfying the boundary condition. This is not true on  $B_n$ , but there is a related result. Let  $\Gamma_n$  be any closed horizontal curve on  $B_h$ , and consider the rate of change of the circulation around  $\Gamma_n$ . We have

$$\begin{aligned} \frac{\partial}{\partial t} \int_{\Gamma_n} \underline{u} \cdot d\underline{r} &= - \int_{\Gamma_n} [2\underline{k} \times \underline{u} + \rho_0^{-1} \nabla p + \rho_0^{-1} \rho \nabla\varphi] \cdot d\underline{r} \\ &= -2 \int_{\Gamma_n} \underline{k} \times \underline{u} \cdot d\underline{r} \end{aligned}$$

since  $\Gamma_n$  is horizontal (so  $\nabla\varphi \cdot d\underline{r} = 0$ ),  $\rho_0$  is constant along  $\Gamma_n$ , and  $\nabla p \cdot d\underline{r}$  is an exact differential. However, on  $\Gamma_n$ ,

$$d\underline{r} = \frac{\nabla\varphi \times \underline{n}}{|\nabla\varphi \times \underline{n}|} d\sigma,$$

$d\sigma$  being arc length, so

$$\underline{k} \times \underline{u} \cdot d\underline{r} = - \frac{\underline{n} \cdot \underline{k}}{|\nabla\varphi \times \underline{n}|} \underline{u} \cdot \nabla\varphi d\sigma = \frac{\underline{n} \cdot \underline{k}}{|\nabla\varphi \times \underline{n}|} \frac{s_t}{s'_0} d\sigma.$$

Thus we obtain

$$\frac{\partial}{\partial t} \int_{\Gamma_n} \underline{u} \cdot d\underline{r} = -2 \frac{\partial}{\partial t} \int_{\Gamma_n} \frac{\underline{n} \cdot \underline{k}}{|\nabla\varphi \times \underline{n}|} \frac{s}{s'_0} d\sigma,$$

so setting

$$C(\Gamma_n) = \int_{\Gamma_n} [\underline{u} \times \nabla\varphi + 2 \frac{s}{s'_0} \underline{k}] \cdot \underline{n} \frac{d\sigma}{|\nabla\varphi \times \underline{n}|} \quad (14)$$

we see that  $C(\Gamma_n)$  is constant in time. We have established our first basic result:

Theorem I. For every solution of (5) - (8) with  $\underline{u} \cdot \underline{n} = 0$  on  $B$ ,

- (a)  $s$  is conserved on the flat spots  $B_h$ , if any, and
- (b)  $C(\Gamma_n)$  (given by (14)) is conserved for each closed horizontal curve  $\Gamma_n$  lying on  $B_n$ .

Our second basic conservation property concerns a quantity related to that appearing in the integrand of (14). One readily shows from (5) and (7) that

$$\frac{\partial}{\partial t} \left\{ \rho_0 \varphi_z^{-1} \left[ \underline{u} \times \nabla\varphi + 2s(s'_0)^{-1} \underline{k} \right] \right\} = -2\rho_0 \underline{u} - \varphi_z^{-1} \nabla p \times \nabla\varphi. \quad (15)$$

Taking the divergence of this and using (6) we get

$$\frac{\partial}{\partial t} \left\{ \nabla \cdot \left[ \rho_0 \varphi_z^{-1} \left( \underline{u} \times \nabla \varphi + 2s(s'_0)^{-1} \underline{k} \right) \right] - 2\rho \right\} = \varphi_z^{-2} \nabla p \cdot \nabla \varphi \times \nabla \varphi_z. \quad (16)$$

Let us define the scalar field  $\Pi$  by

$$\Pi = \rho_0^{-1} \nabla \cdot \left[ \rho_0 \varphi_z^{-1} \left( \underline{u} \times \nabla \varphi + 2(s'_0)^{-1} s \underline{k} \right) \right] - 2\rho_0^{-1} \rho. \quad (17)$$

We call  $\Pi$  the "potential vorticity" of the flow. Throughout any geostrophically free region we see at once from (16) that  $\Pi$  is conserved pointwise. In non-free regions this is not true, but a somewhat weaker kind of conservation does apply. First consider a geostrophically guided region, and let  $\Gamma$  be any closed geostrophic curve in it (or on its boundary). We define the "mean potential vorticity for  $\Gamma$ " by the formula:

$$\bar{\Pi}(\Gamma) = \int_{\Gamma} \varphi_z^2 |\nabla \varphi \times \nabla \varphi_z|^{-1} \Pi \, d\sigma. \quad (18)$$

Along  $\Gamma$  we have  $d\underline{r} = |\nabla \varphi \times \nabla \varphi_z|^{-1} \nabla \varphi \times \nabla \varphi_z \, d\sigma$ , so (16) shows that

$$\begin{aligned} \frac{d}{dt} \bar{\Pi}(\Gamma) &= \int_{\Gamma} \rho_0^{-1} |\nabla \varphi \times \nabla \varphi_z|^{-1} \nabla p \cdot \nabla \varphi \times \nabla \varphi_z \, d\sigma \\ &= \int_{\Gamma} \rho_0^{-1} \nabla p \cdot d\underline{r} = 0, \end{aligned}$$

since  $\rho_0$  is constant along  $\Gamma$  and  $\nabla p \cdot d\underline{r}$  is exact. This kind

of a result does not apply in blocked regions, since there the geostrophic curves are not closed. For a blocked region we consider any horizontal surface  $\Sigma$  lying in the blocked region, and let  $\Gamma$  be the horizontal curve on  $B$  which is the boundary of  $\Sigma$ ;  $\Sigma$  and  $\Gamma$  are characterized by some value of the potential, say  $\varphi_0$ . Then we construct the quantify

$$Q(\varphi_0) = \int_{\Sigma} \Pi |\nabla \varphi|^{-1} dS - \int_{\Gamma} \varphi_z^{-1} |\nabla \varphi \times \underline{n}|^{-1} (\underline{u} \times \nabla \varphi + 2s(s')^{-1} \underline{k}) \underline{n} d\sigma \quad (19)$$

where the orientations of  $\Gamma$  and  $\Sigma$  are related as usual in Stokes' theorem, i.e. if  $|\nabla \varphi|^{-1} \nabla \varphi$  is regarded as the unit normal to  $\Sigma$ , on  $\Gamma$  we have  $d\underline{r} = |\nabla \varphi \times \underline{n}|^{-1} \nabla \varphi \times \underline{n} d\sigma$ . Now we have on  $\Gamma$ :

$$\begin{aligned} & (\underline{u}_t \times \nabla \varphi + 2s_t(s'_0)^{-1} \underline{k}) \cdot \underline{n} |\nabla \varphi \times \underline{n}|^{-1} d\sigma \\ &= \left[ (-2\underline{k} \times \underline{u} - \rho_0^{-1} \nabla p) \times \nabla \varphi - 2\underline{u} \cdot \nabla \varphi \underline{k} \right] \cdot \underline{n} |\nabla \varphi \times \underline{n}|^{-1} d\sigma \\ &= \left[ -2\varphi_z \underline{u} - \rho_0^{-1} \nabla p \times \nabla \varphi \right] \cdot \underline{n} |\nabla \varphi \times \underline{n}|^{-1} d\sigma = -\rho_0^{-1} \nabla p \cdot d\underline{r}, \quad \text{so} \end{aligned}$$

$$\begin{aligned} \frac{dQ}{dt} &= \int_{\Sigma} \rho_0^{-1} \varphi_z^{-2} \nabla p \cdot \nabla \varphi \times \nabla \varphi_z |\nabla \varphi|^{-1} dS + \int_{\Gamma} \varphi_z^{-1} \rho_0^{-1} \nabla p \cdot d\underline{r} \\ &= \int_{\Sigma} -\rho_0^{-1} \nabla \times (\varphi_z^{-1} \nabla p) \cdot \nabla \varphi |\nabla \varphi|^{-1} dS + \int_{\Gamma} \varphi_z^{-1} \rho_0 \nabla p \cdot d\underline{r} = 0 \end{aligned}$$

by Stokes' Theorem, since  $\rho_0$  is constant on  $\Sigma$ . This finishes the proof of

Theorem II. For every solution of (5) - (8) with  $\underline{u} \cdot \underline{n} = 0$  on  $B$ ,

- (a)  $\Pi$  (given by (17)) is conserved at each point of a geostrophically free region,
- (b)  $\bar{\Pi}$  (given by (18)) is conserved for each geostrophic curve in a guided region, and
- (c)  $Q$  (given by (19)) is conserved for each horizontal surface in a blocked region.

5. The initial-value problem. Suppose  $\underline{u}(t)$ ,  $\rho(t)$ ,  $s(t)$ ,  $p(t)$  is the solution to the initial value problem for some given set of initial conditions. We shall define the geostrophic part of  $\underline{u}$ , etc., (called  $\underline{u}_g$ , etc.) to be the projection, orthogonal in the sense of the inner product (11), of  $\underline{u}$ , etc. onto the space of geostrophic flows. Of course we really have such a projection for each value of  $t$ , but these are in fact always the same. This would be immediately obvious if we knew the existence and completeness of normal mode solutions in terms of which we could represent  $\underline{u}(t)$  etc., for all oscillatory normal modes are orthogonal to the steady ones, which are just the geostrophic flows. But even without the existence of a complete set of normal modes, the projection is easily seen to be independent of time; this follows for instance from the equivalent definition of  $\underline{u}_g$  etc. as that geostrophic flow which minimizes the energy of the difference flow  $\underline{u} - \underline{u}_g$ , etc., (at a given time  $t$ ). If  $\underline{u}_{g_1}$  is the projection at  $t_1$  and  $\underline{u}_{g_2}$  that at  $t_2$  then the energies

of the difference flows  $\underline{u} - \underline{u}_{g_1}$  and  $\underline{u} - \underline{u}_{g_2}$  must be constant in time, hence equal, since one is minimal at  $t_1$  and the other at  $t_2$ . The identity of  $\underline{u}_{g_1}$  and  $\underline{u}_{g_2}$  then follows from the uniqueness of the projection implied by the positive definiteness of the energy. Since the geostrophic part is independent of time, it can, in principle at least, be determined from the initial data - it is our main purpose to describe explicitly how this may be done. This problem is not altogether trivial because the space of geostrophic flows is relatively extensive - it is not simply a matter of calculating a coefficient, as it would be if there were only one geostrophic flow.

It is a reasonable conjecture that the geostrophic part might also be the time average over  $0 \leq t < \infty$  of the solution to the initial value problem. This is probably essentially true; for example if one assumes that  $\underline{u}(t)$  etc. remain bounded and that the time average exists, one easily checks that the time average satisfies the geostrophic equations. Then the time average of  $\underline{u} - \underline{u}_g$  etc. would also be a geostrophic flow, and one that is orthogonal to all geostrophic flows (assuming interchangeability of spacial and temporal integrations), in particular to itself. Thus the time average of  $\underline{u}$  would in fact be  $\underline{u}_g$ . It should be remarked however that this argument, even if it could be fully justified, would only prove that  $\underline{u}_g$  was the time average almost everywhere. It seems in fact to be the case that sometimes the time average is not exactly the same as  $\underline{u}_g$ , but exhibits certain discontinuities on parts of the boundary. We shall see later how this comes about.

To determine  $\frac{\underline{u}}{g}$  we first need a formula for the inner product of an arbitrary time dependent flow  $\underline{u}, p, \rho, s$  with an arbitrary geostrophic flow  $\hat{\underline{u}}, \hat{p}, \hat{\rho}, \hat{s}$ ; the latter is given in terms of its pressure field  $\hat{p}$  by the formulas of §3, and we want to express the inner product also in terms of  $\hat{p}$ . We have, using the real form of (11) (the complex form is only needed in connection with oscillatory normal modes):

$$\begin{aligned}
 \langle \hat{\underline{u}}, \hat{p}, \hat{\rho}, \hat{s} | \underline{u}, p, \rho, s \rangle &= \int_R \left[ \rho_0 \hat{\underline{u}} \cdot \underline{u} + \rho_0^{-1} f_p \hat{p} p - \rho_0^{-1} (s'_0)^{-1} \hat{s} s \right] dV \\
 &= \int_R \left[ \frac{1}{2} \varphi_z^{-1} \nabla \varphi \times \nabla \hat{p} \cdot \underline{u} + \rho_0^{-1} f_p \hat{p} p + (s'_0)^{-1} s \left( \varphi_z^{-1} \hat{p}_z + f_p \hat{p} \right) \right] dV \\
 &= \int_R \left[ \varphi_z^{-1} \left( \frac{1}{2} \underline{u} \times \nabla \varphi + (s'_0)^{-1} s \underline{k} \right) \cdot \nabla \hat{p} + f_p \hat{p} \left( \rho_0^{-1} p + (s'_0)^{-1} s \right) \right] dV \\
 &= \int_B \frac{1}{2} \varphi_z^{-1} \hat{p} \underline{n} \cdot \left[ \underline{u} \times \nabla \varphi + 2(s'_0)^{-1} s \underline{k} \right] dS + \int_R f_p \hat{p} \left( \rho_0^{-1} p + (s'_0)^{-1} s \right) dV \\
 &\quad - \int_R \frac{1}{2} \hat{p} \nabla \cdot \left[ \varphi_z^{-1} \left( \underline{u} \times \nabla \varphi + 2(s'_0)^{-1} s \underline{k} \right) \right] dV.
 \end{aligned}$$

But

$$\begin{aligned}
 \nabla \cdot \left[ \varphi_z^{-1} \left( \underline{u} \times \nabla \varphi + 2(s'_0)^{-1} s \underline{k} \right) \right] &= \rho_0^{-1} \nabla \cdot \left[ \rho_0 \varphi_z^{-1} \left( \underline{u} \times \nabla \varphi + 2(s'_0)^{-1} s \underline{k} \right) \right] \\
 &= 2\rho_0^{-1} \rho'_0(\varphi)(s'_0)^{-1} s = \Pi + 2\rho_0^{-1} \rho - 2\rho_0^{-1} (f_s s'_0 - f_p \rho_0)(s'_0)^{-1} s \\
 &= \Pi + 2\rho_0^{-1} f_p (p + \rho_0 (s'_0)^{-1} s).
 \end{aligned}$$

Thus we obtain:

$$\begin{aligned}
\langle \hat{\underline{u}}, \hat{\underline{p}}, \hat{\underline{\rho}}, \hat{\underline{s}} | \underline{u}, p, \rho, s \rangle = & -\frac{1}{2} \int_R \hat{p} \Pi \, dV + \int_{B_h} \varphi_z^{-1} \hat{p}(s'_0)^{-1} \underline{s} \underline{n} \cdot \underline{k} dS \\
& + \frac{1}{2} \int_{B_n} \varphi_z^{-1} \hat{p} \underline{n} \cdot [\underline{u} \times \nabla \varphi + 2(s'_0)^{-1} \underline{s} \underline{k}] dS. \quad (20)
\end{aligned}$$

Now  $\hat{p}$  is the pressure field of a geostrophic flow, so it must be constant along the horizontal curves  $\Gamma_n(\varphi)$  on  $B_n$ . We imagine the last surface integral in (20) to be calculated by dividing  $B_n$  up into infinitesimal strips between adjacent curves  $\Gamma_n$ ; one easily sees from the geometry that in this case we have  $dS = |\nabla \varphi \times \underline{n}|^{-1} d\varphi \, d\sigma$ , so

$$\begin{aligned}
& \int_{B_n} \varphi_z^{-1} \hat{p} \underline{n} \cdot [\underline{u} \times \nabla \varphi + 2(s'_0)^{-1} \underline{s} \underline{k}] dS \\
& = \int_{B_n} \hat{p}(\varphi) d\varphi \int_{\Gamma_n(\varphi)} \varphi_z^{-1} |\nabla \varphi \times \underline{n}|^{-1} \underline{n} \cdot (\underline{u} \times \nabla \varphi + 2(s'_0)^{-1} \underline{s} \underline{k}) d\sigma \quad (21)
\end{aligned}$$

We now use these formulas to characterize those flows which are orthogonal to all geostrophic flows, by requiring the right hand side of (20) to be zero for an arbitrary geostrophic pressure field  $\hat{p}$ . Let us suppose the region  $R$  to be divided into geostrophically free, guided, and blocked parts,  $R_F$ ,  $R_G$  and  $R_B$ , and denote the portions of  $B_h$  and  $B_n$  which are associated with each such part by similar subscripts. First we may choose  $\hat{p}$  to be zero on the boundary and in  $R_G$  and  $R_B$ , but otherwise arbitrary in  $R_F$ . (20) then shows that for  $\underline{u}$ , etc., to be orthogonal to all geostrophic flows it must have zero potential vorticity throughout  $R_F$ . Next we may choose a  $\hat{p}$  which is zero in  $R_G$  and  $R_B$  (as we

have just seen the volume integral then vanishes since  $\Pi$  must be zero in  $R_F$  and which is zero on  $B_n$ ,  $B_{nG}$  and  $B_{nB}$ , but which is arbitrary on  $B_{nF}$ . We thus see that it is also necessary that  $s$  should vanish (pointwise) on  $B_{nF}$ . Now we cannot choose a geostrophic pressure field which is arbitrary on  $B_{nF}$ , for there it must be constant along horizontal curves, but its value can vary arbitrarily from one horizontal curve to the next. Using the form (21) for the surface integral over  $B_{nF}$  we see then, noting that  $\varphi_z$  is constant along horizontal curves on  $B_{nF}$ , that it is also necessary that the quantity  $C$  defined by (14) should be zero on every horizontal curve on  $B_{nF}$ . Next we may choose a  $\hat{p}$  which is zero on the boundary and in  $R_B$  but is arbitrary in  $R_G$ , to the extent that is possible, i.e. an arbitrary function of geostrophic curves in  $R_G$ . Since  $\hat{p}$  is not completely arbitrary we cannot conclude that  $\Pi$  must vanish in  $R_G$ , but if we think of the volume integral over  $R_G$  as being evaluated by integrating over infinitesimal tubes around the geostrophic curves, such a tube being given for instance by  $\varphi_1 \leq \varphi \leq \varphi_1 + d\varphi$  and  $\varphi_{z1} \leq \varphi_z \leq \varphi_{z1} + d\varphi_z$ , one sees that  $dV = |\nabla\varphi \times \nabla\varphi_z|^{-1} d\varphi d\varphi_z d\sigma$ , and the vanishing of the volume integral for arbitrary  $\hat{p}(\varphi, \varphi_z)$  thus implies that the mean potential vorticity  $\bar{\Pi}(\Gamma)$  defined by (18) must vanish for every closed geostrophic curve in  $R_G$ . As above one then shows that  $C(\Gamma_n)$  must also vanish for  $\Gamma_n$ 's on  $B_{nG}$ , since  $\varphi_z$  is constant along these curves, the geostrophic curves in  $R_G$  not crossing the boundary. However one cannot conclude that  $s = 0$  pointwise on  $B_{nG}$ , for  $\hat{p}$  must be constant along the geostrophic curves there.

Dividing  $B_{hG}$  into strips between adjacent geostrophic curves we find  $dS = |\nabla\varphi \times \nabla\varphi_z|^{-1} |\nabla\varphi| d\varphi_z d\sigma$  and thus since  $\underline{n} \cdot \underline{k} = \pm \varphi_z |\nabla\varphi|^{-1}$  on  $B_h$ , we see that for each geostrophic curve  $\Gamma_h$  on  $B_{hG}$  the quantity

$$D(\Gamma_h) = \int_{\Gamma_h} |\nabla\varphi \times \nabla\varphi_z|^{-1} d\sigma \quad (21)$$

must vanish. Passing now to the blocked parts, if any, we note that here the arbitrariness of  $\hat{p}$  is so restricted that we can no longer consider the volume integral and the surface integral over  $B_{nB}$  separately. Evaluating the former by integrating over horizontal sheets between adjacent surfaces of constant  $\varphi$  and the latter by using the corresponding horizontal strips on  $B_{nB}$  we find that the necessary condition is the vanishing of the quantity  $Q(\Sigma)$  defined by (19), for each horizontal surface  $\Sigma$  in  $R_B$ . Finally on  $B_{hB}$ , since  $\underline{n} \cdot \underline{k} = \pm |\nabla\varphi|^{-1} \varphi_z$  there, we must have the quantity

$$E = \int_{B_{hB}} |\nabla\varphi|^{-1} dS \quad (22)$$

zero as well. It should perhaps be mentioned explicitly that in case some of the surfaces  $\Sigma$  or  $B_h$  consist of several disconnected pieces, these conditions must hold separately for each connected component. Summarizing these results we have

Theorem III. In order that a flow  $\underline{u}$ ,  $p$ ,  $\rho$ ,  $s$  should be orthogonal to all geostrophic flows it is necessary (and sufficient) that each of the following should vanish:

- (a) The potential vorticity  $\Pi$  (eq. (17)), at each point of  $R_F$ ;
- (b)  $\bar{\Pi}(\Gamma)$  (eq. (18)), for each geostrophic curve  $\Gamma$  in  $R_G$ ;
- (c)  $Q(\Sigma)$  (eq. (19)), for each horizontal surface in  $R_B$ ;
- (d)  $C(\Gamma_n)$  (eq. (14)), for each horizontal curve on  $B_{nF}$  or  $B_{nG}$ ;
- (e)  $s$ , on  $B_{hF}$ ;
- (f)  $D(\Gamma_h)$  (eq. (21)), for each geostrophic curve  $\Gamma_h$  on  $B_{hG}$ ;
- (g)  $E$  (eq. (22)), for each connected component of  $B_{hB}$ .

Notice that all of these quantities are constant in time, because of Theorems I and II. They must of course be, because of the temporal invariance of the inner product which assures us that if  $\underline{u}$  etc. is once orthogonal to all geostrophic flows it will always be so; note however that Theorem I is slightly stronger since it asserts the constancy of  $s$  pointwise on all of  $B_h$ , without the averaging implied in the quantities  $D$  and  $E$ .

Theorem III now allows us to characterize completely the geostrophic part of the solution to the initial value problem. For  $\underline{u}_g$  etc. to be the projection of  $\underline{u}$  etc. onto the geostrophic flows simply means that the difference flow  $\underline{u} - \underline{u}_g$  etc. is orthogonal to all geostrophic flows, hence the quantities (a) ... (g) mentioned in Theorem III must all vanish when calculated for this difference flow. Thus  $\underline{u}_g$  etc. must be a geostrophic flow for which the quantities (a) ... (g) are all the same as they are for the initial data. The geostrophic part is uniquely determined by these conditions, for the difference of two geostrophic flows both of which have the

same values of (a) ... (g) would be a geostrophic flow with zero values of these quantities, hence orthogonal to itself and so zero. This completes

Theorem IV. The geostrophic part of the solution to the initial value problem is that (unique) geostrophic flow which has the same values of the quantities mentioned in Theorem III as the initial flow has.

The actual computation of  $\frac{u}{g}$  may not be easy, but it is easier than the full initial value problem. For example if the entire region is geostrophically free, the condition  $\Pi = \Pi_0$  gives an elliptic partial differential equation in 3 variables for  $p_g$ , for which in addition we have the boundary conditions that  $p_g$  is constant on horizontal curves on  $B_h$ ,  $C = C_0$  on these curves, and  $s = s_0$  on  $B_h$ . We do not have a proof of the existence of a solution to this problem (as usual, uniqueness is much easier) but we shall see an example in the next section which at least suggests that such existence is to be expected.

The unique determination of the geostrophic part by the initial values of the quantities (a) ... (g) indicates that we may describe them as "the features of the initial data which are carried by the geostrophic part." It is natural to expect that the geostrophic part would carry all aspects of the initial data which are constant in time - yet this is not quite so, at least when guided or blocked regions are present, for Theorem I shows for instance that  $s$  is

pointwise constant on  $B_n$ , yet the geostrophic part can only carry certain averages of  $s$  on  $B_{nG}$  and  $B_{nB}$ ; similarly  $C(\Gamma_n)$  is conserved for all  $\Gamma_n$ 's on  $B_n$ , but the geostrophic part may not be able to carry all of  $C$  on  $B_{nB}$ , though it does on the rest of  $B_n$ . It seems however that this somewhat peculiar situation occurs only at the boundary; at any rate simple examples which can be explicitly solved and in which similar phenomena occur indicate this, and we have seen above that with fairly plausible assumptions one can show the identity of the geostrophic part and the time average almost everywhere.

5. Examples and remarks. Let us consider first the situation typical of most laboratory experiments where the gravitational field is that of the earth; to have it constant in the rotating system we must assume that the vectors  $\underline{\Omega}$  and  $\underline{g}$  are aligned, so that the dimensionless potential function has the form

$$\varphi = -\frac{1}{2}(x^2 + y^2) + \gamma z \quad (23)$$

where  $\gamma = g/(\Omega^2)$ . In this case the entire region is geostrophically free. (If  $\underline{\Omega}$  and  $\underline{g}$  are not parallel, we will not of course have any steady basic state; this is indeed one of the difficulties in doing experiments with rotating stratified fluids - the alignment must be quite precise or it will be impossible to produce a sufficiently stationary basic state.) Most laboratory experiments are also done with fluids for which some sort of Boussinesq approximation is appropriate.

In our non-dissipative case we may take this to mean that variations in density due to pressure alone are negligible ( $f_p = 0$ ; this eliminates compression waves) and variations in density due to heating (or salt;  $s$  may be taken to be salinity rather than entropy if the stratification is produced that way) are small compared to the average density. A convenient simple model is obtained by taking the basic stratification to be such that  $s'_0$  is constant, and assuming that also  $f_s = -\alpha$  is constant. With a suitable choice of the basic scales of density and entropy we may take  $\rho_0 \cong 1$ ,  $s'_0 = \gamma^{-1}$ , and  $\rho = -\alpha s$  as the linearized equation of state.  $\rho$  may then be eliminated from the problem, and the basic equations of this model become

$$\underline{u}_t + 2\underline{k} \times \underline{u} + \nabla p - \alpha \gamma s (\underline{k} - \gamma^{-1}(\underline{x}\underline{i} + \underline{y}\underline{j})) = 0 \quad (24)$$

$$\nabla \cdot \underline{u} = 0 \quad (25)$$

$$s_t + \underline{u} \cdot (\underline{k} - \gamma^{-1}(\underline{x}\underline{i} + \underline{y}\underline{j})) = 0. \quad (26)$$

This may be still further simplified by taking  $\gamma$  to be large, supposing that  $\alpha\gamma \equiv N^2$  is of order unity. This makes the horizontal surfaces perpendicular to  $\underline{k}$ . In this case the energy is

$$\mathcal{E} = \frac{1}{2} \int_R [|\underline{u}|^2 + N^2 s^2] dV \quad (27)$$

and the potential vorticity is

$$\Pi = \nabla \cdot [\underline{u} \times \underline{k} + 2s\underline{k}] = \underline{k} \cdot \nabla \times \underline{u} + 2s_z. \quad (28)$$

The geostrophic flows are given by

$$\underline{u} = \frac{1}{2} \underline{k} \times \nabla p \quad (29)$$

and

$$s = N^{-2} p_z \quad (30)$$

for any function  $p$  constant along horizontal curves on  $B_n$ . The potential vorticity of a geostrophic flow is thus given by

$$\Pi_g = \frac{1}{2} \left( \nabla_H^2 p + 4N^{-2} p_{zz} \right) \quad (31)$$

where

$$\nabla_H = \nabla - \underline{k} \frac{\partial}{\partial z}.$$

The quantity  $C$  of (14) is

$$C(\Gamma_n) = \int_{\Gamma_n} [\underline{u} \times \underline{k} + 2s\underline{k}] \cdot \underline{n} |\underline{k} \times \underline{n}|^{-1} d\sigma \quad (32)$$

which for a geostrophic flow becomes

$$C = \frac{1}{2} \int_{\Gamma_n} [\nabla_H^2 p + 4N^{-2} p_z \underline{k}] \cdot \underline{n} |\underline{k} \times \underline{n}|^{-1} d\sigma. \quad (33)$$

A specific problem for which explicit calculations can readily be made is obtained with this model by taking the region  $R$  to be the cylinder:  $0 \leq x^2 + y^2 \leq a^2$ ,  $0 \leq z \leq 1$ . To compute the geostrophic part of the solution to the initial value problem we compute

$\Pi_0 = \mathbf{k} \cdot \nabla \times \mathbf{u}_0 + 2s_{0z}$  from the initial data, and also (going to cylindrical coordinates)

$$C_0(z) = \int_0^{2\pi} \mathbf{u}_0(a, \theta, z) \cdot \mathbf{e}_1 a d\theta.$$

The problem for the geostrophic pressure field  $p$  is thus

$$\nabla_H^2 p + 4N^{-2} p_{zz} = 2\Pi_0 \quad (34)$$

with boundary conditions

$$p_z = N^2 s_0 \quad \text{on } z = 0, 1 \quad (35)$$

$$p = \text{const. on } r = a \quad \text{for each fixed } z \quad (36)$$

$$\int_0^{2\pi} \frac{\partial p}{\partial r}(a, z, \theta) a d\theta = 2C_0(z). \quad (37)$$

The somewhat unusual boundary conditions (36) and (37) on the non-horizontal boundary become familiar ones when  $p$  is represented as a Fourier series in  $\theta$ :

$$p = \sum_{-\infty}^{\infty} p^{(m)}(r, z) e^{im\theta};$$

(36) thus is

$$p^{(m)}(a, z) = 0 \quad m \neq 0$$

and (37) gives

$$\frac{\partial p^{(0)}}{\partial r}(a, z) = \frac{a}{\pi} C_0(z).$$

To verify the existence of a solution to the Neumann problem for  $p^{(0)}$  it is only necessary to check that the source strength  $2\pi^{(0)}$  in (34) is consistent with the net flux out through the boundaries given by (35) and (37); this is readily done. The solution then can be computed for example by expanding  $p^{(0)} - ra/\pi C_0(z)$  in a series of the form  $\sum a_n(z) J_0(k_n r/a)$ , where the  $k_n$  are roots of  $J'_0$ , and  $p^{(m)}$  in series  $\sum a_n^{(m)}(z) J_{|m|}(k_n^{(m)} r/a)$ , where the  $k_n^{(m)}$  are roots of  $J_{|m|}$ . An alternative representation is obtainable by removing the inhomogeneity in the boundary condition (35) instead of that in (37), and using Fourier cosine series in  $z$ .

With this simple model it is in fact possible to find explicitly all the oscillatory normal modes as well, and so solve the full initial value problem. Although this brings out a number of interesting features, we do not do this here, our main interest in this paper being in the geostrophic part.

Interesting examples which can be solved explicitly to the same extent as our first one do not seem to be so easily found in the geostrophically guided case, but without attempting a full discussion

we may consider the following. We take the same sort of Boussinesq fluid as above, but now the region will be a spherical shell given in dimensionless form (spherical coordinates) by  $a \leq r \leq a + 1$ . For the geopotential we shall take  $\Phi = -\gamma a^2/r$  so that "gravity" points radially inward, and as above we assume uniform stratification, with scaling such that  $s'_0 = \gamma^{-1}$ . We also suppose  $\gamma$  is large and  $\alpha$  small so that  $N^2 \equiv \alpha\gamma$  is of order unity. Then we find  $\nabla\Phi = \underline{r}_1 \gamma a^2/r^2$ , where  $\underline{r}_1$  is the radial unit vector; and  $\Phi_z = \gamma a^2 z/r^3$ . Thus  $\nabla\Phi \times \nabla\Phi_z = \underline{r}_1 \times \underline{k} \gamma^2 a^4/r^5$ , and the geostrophic curves are latitude circles on the spheres of constant  $r$ . The potential vorticity is found to be

$$\Pi = \nabla \cdot \left[ \sec \theta \left( \underline{u} \times \underline{r}_1 + 2r^2 s/a^2 \underline{k} \right) \right], \quad (38)$$

$\theta$  being the polar angle. The geostrophic flows are given by

$$\underline{u}_g = \frac{1}{2} \sec \theta \underline{r}_1 \times \nabla p \quad (39)$$

and

$$s_g = r^2 a^{-2} N^{-2} \sec \theta p_z = r^2 a^{-2} N^{-2} (p_r - \tan \theta p_\theta/r) \quad (40)$$

where  $p = p(r, \theta)$  is constant on the geostrophic curves, but is otherwise arbitrary since the boundaries  $r = a, a + 1$  are entirely horizontal. Computing the potential vorticity for a geostrophic flow one obtains:

$$\begin{aligned}
\Pi_g &= 2N^{-2}r^{-2} \frac{\partial}{\partial r} \left\{ r^6 a^{-4} \left( p_r - \tan \theta p_{\theta}/r \right) \right\} \\
&\quad + \frac{1}{2r \sin \theta} \frac{\partial}{\partial \theta} \left\{ \sin \theta \left[ \left( \sec^2 \theta + 4N^{-2}r^4 a^{-4} \tan^2 \theta \right) p_{\theta}/r \right. \right. \\
&\quad \left. \left. - 4N^{-2}r^4 a^{-4} \tan \theta p_r \right] \right\} \\
&= \frac{2a^2}{N^2 r^2 \sin \theta} \left[ \frac{\partial}{\partial r}, \frac{1}{r} \frac{\partial}{\partial \theta} \right] r^6 a^{-6} \sin \theta \begin{bmatrix} 1 & -\tan \theta \\ -\tan \theta \tan^2 \theta + \frac{N^2 a^4}{4r^4} \sec^2 \theta \end{bmatrix} \begin{bmatrix} \frac{\partial p}{\partial r} \\ \frac{1}{r} \frac{\partial p}{\partial \theta} \end{bmatrix}
\end{aligned} \tag{41}$$

Since the weight factor in (18) is independent of azimuth, as is  $\Pi_g$ , the basic equation  $\bar{\Pi}_g = \bar{\Pi}_0$  is equivalent to setting  $\Pi_g$  equal to the average over azimuth of the initial potential vorticity field. Similarly the boundary conditions on  $p$ ,  $D_g = D_0$  (cf. (21)), are equivalent to taking  $r^2 a^{-2} N^{-2} (p_r - \tan \theta p_{\theta}/r)$  to be the azimuthal average of the initial entropy on  $r = a$  and  $a + 1$ . The explicit calculation of the geostrophic part of the flow produced by given initial conditions appears to be a fairly difficult problem in general with this model, but by carefully choosing the initial conditions it is reasonably easy to give some simple examples. One class of such examples is obtained with initial conditions of the form:

$$u_0 = 0, \quad s_0 = A(r) \cos^3 \theta + B(r) \cos \theta \tag{42}$$

(with some restrictions on  $A$  and  $B$  to be mentioned in a moment),  
or any initial conditions whose azimuthal averages are of this form.  
In this case the potential vorticity is found from (38) to be

$$\Pi_0 = 2r^2 a^{-2} A'(r) \cos^3 \theta + 2a^{-2} [(r^2 B)'] + 2rA] \cos \theta. \quad (43)$$

A little calculation shows that a geostrophic flow determined by  
 $p = f(r) \cos^3 \theta$  will give this same potential vorticity field provided  
that  $A$  and  $B$  are related to  $f$  by

$$A = a^{-2} N^{-2} r^5 (f/r^3)', \quad (44)$$

and

$$B = 3a^{-2} N^{-2} r f - (3a^2/2r^2) \int_a^r f(r_1) r_1^{-2} dr_1. \quad (45)$$

Such a geostrophic flow has entropy

$$s_g = a^{-2} N^{-2} \left[ r^5 (f/r^3)' \cos^3 \theta + 3rf \cos \theta \right] \quad (46)$$

and thus as far as the  $\cos^3 \theta$  term is concerned the conditions that  
 $s_g = s_0$  on  $r = a$  and  $a + 1$  are satisfied for any  $f$ ; but the  
 $\cos \theta$  term shows that we also must require that

$$\int_a^{a+1} f(r) r^{-2} dr = 0.$$

Thus to construct an example we need only take some  $f$  satisfying this condition, and calculate  $A$  and  $B$  from (45) and (46).  $f \cos^3 \theta$  will then be the pressure field of the geostrophic part of the flow that evolves from the initial state (42). Its velocity field is, from (39),

$$\frac{u}{g} = -\frac{3}{4} \sin 2\theta (f/r) \underline{\phi}_1 \quad (47)$$

where  $\underline{\phi}_1$  is the unit azimuthal vector; the entropy (46) may also be written in the form:

$$s_g = s_0 + \frac{3}{2} a^2 r^{-2} \cos \theta \int_a^r f(r_1) r_1^{-2} dr_1. \quad (48)$$

A particularly simple explicit example is obtained by taking  $f = r^2(r - a - \frac{1}{2})/a^3$ ; in this case, if  $N^2$  is moderate and  $a$  is rather large (thin shell) the initial entropy field is not too different from  $\cos^3 \theta / (aN^2)$ , and is thus a warming of the northern and cooling of the southern hemispheres. The geostrophic velocity field (47) is then vertically sheared, being from west to east at low altitudes and from east to west at high altitudes in the northern hemisphere, and reversed in the southern.

It is of interest to compare our general results with their analogues for a rotating homogeneous fluid, for the relationship between the two is a little more subtle than one might perhaps naively expect. Of course one would anticipate that "turning on" a little stratification (with a gravitational field) would be a small

perturbation, at least for a finite time, on the full initial value problem, but this is not necessarily the case when attention is focussed on the steady part of the solution.. Indeed steady flows in the linear theory of rotating homogeneous fluids must be independent of  $z$  (Taylor-Proudman Theorem), and this is not at all necessary for steady stratified flows; on the other hand steady stratified flows must be along geopotential surfaces, which is not necessary in the homogeneous case where the geopotential (in the absence of free surfaces) plays no role at all. What is happening here is of course that some of the oscillatory normal modes in the stratified case have frequencies which tend to zero as the stratification disappears and so show up as geostrophic motions when the stratification is turned off. Likewise, some initial conditions which would excite time dependent motion (probably largely at low frequencies) in the homogeneous case, may in fact lead to steady flows when a little stratification is put in. Though the situation is really more complex, it is a little like the apparent discontinuity in the nature of geostrophic homogeneous flows when the bottom of a container of constant height is slightly deformed: the class of geostrophic flows is greatly curtailed, and new low frequency oscillatory modes (Rossby waves) appear.

The homogeneous analogues of our results can of course be obtained by the same methods - the difference comes only in the first step, the

description of the geostrophic flows themselves. Although these results are already fairly well known (they are essentially given, for instance, in H. P. Greenspan's monograph "The theory of rotating fluids," Cambridge, 1968) it is perhaps worth recalling them briefly to emphasize that the analogy with the stratified case is in fact quite close. To describe the homogeneous geostrophic flows one must also distinguish free, guided, and blocked regions. The free regions are those of constant height (measured parallel to the rotation axis); the guided regions are those not of constant height, but in which the contours of constant height ("geostrophic curves") do not cross the boundaries; and the blocked regions are those in which they do. The geostrophic flows in free regions are independent of  $z$  and parallel to the top and bottom, but otherwise arbitrary; they are essentially characterized by the vertical vorticity component, a quantity which, when vertically averaged, is conserved for any motion in a free region. The geostrophic flows in guided regions are independent of  $z$  and along geostrophic curves; they are characterized by their "mean circulation," the vertically averaged circulation around the closed geostrophic contours, another conserved quantity, in guided regions. The only geostrophic "flow" in a blocked region is zero. The steady part of the solution to an initial value problem "carries" the vertically averaged vertical vorticity of the initial data in free regions and the mean circulation of the initial data in guided regions, just as in the stratified case it carries the potential vorticity and the mean potential vorticity. In over-all structure the stratified geostrophic flow problem is really very similar to the homogeneous case,

even though the latter is not readily derived from the former as a limiting case.

What is the relationship between our function  $\Pi$  and the familiar potential vorticity of non-linear shallow water theory? The latter, the (absolute) vorticity divided by the local depth of the fluid layer, is a quantity which is conserved following particles, not conserved at a particular place like  $\Pi$  (or  $\bar{\Pi}$ ), so a direct analogy is not to be expected; but they are in fact closely related. The true generalization of the potential vorticity of shallow water theory is given by Ertel's Theorem, which (in a non-rotating system) asserts the constancy following particles of the quantity  $\rho^{-1} \nabla \times \underline{u} \cdot \nabla s$ . This holds for an inviscid compressible fluid subject to a conservative body force per unit mass, with  $s$  any scalar field (in particular the entropy, if there is no heat conduction) which is itself constant following particles and in addition, if  $\nabla p \times \nabla \rho \neq 0$ , is a function of  $p$  and  $\rho$ . Ertel's Theorem can be readily proved from the Helmholtz vorticity theorem (which is true under these hypotheses) on using the above properties of  $s$ . The conservation of potential vorticity in shallow water theory follows from Ertel's Theorem, since shallow water theory is mathematically analogous to two dimensional compressible flow of an isentropic gas with  $\gamma = 2$ ; in this analogy the "gas density" is the mass per unit area of the fluid layer, hence proportional to the local depth. The gas motion being two dimensional and isentropic, we may take  $s = z$ , and the conservation of potential vorticity follows. Thus the quantity  $\rho^{-1} \nabla \times \underline{u} \cdot \nabla s$ , or (in a rotating system)  $\rho^{-1} (2\Omega \underline{k} + \nabla \times \underline{u}) \cdot \nabla s$  might appropriately be called "potential

vorticity" for compressible stratified flow. In our linearized theory this is given in dimensionless form by

$$\begin{aligned} & \rho_0^{-1}(1 + \varepsilon \rho/\rho_0)^{-1}(2\underline{k} + \varepsilon \nabla \times \underline{u}) \cdot (s'_0 \nabla \varphi + \varepsilon \nabla s) \\ &= s'_0 \rho_0^{-1} \varphi_z \left[ 2 + \varepsilon \left( \varphi_z^{-1} \nabla \times \underline{u} \cdot \nabla \varphi + 2(s'_0)^{-1} s_z - 2\rho_0^{-1} \rho \right) \right]. \end{aligned} \quad (49)$$

The coefficient of  $\varepsilon$  in this expression is closely related to  $\Pi$ , and our basic formula (16) from which the conservation of  $\Pi$  (or  $\bar{\Pi}$ ) was derived is an immediate consequence of Ertel's Theorem in the linearized case. Indeed we have from (17) that

$$\begin{aligned} \Pi = & \varphi_z^{-1} \nabla \times \underline{u} \cdot \nabla \varphi + 2(s'_0)^{-1} s_z - 2\rho_0^{-1} \rho \\ & + \nabla(\varphi_z^{-1}) \cdot \underline{u} \times \nabla \varphi + 2\rho_0^{-1} s \frac{\partial}{\partial z} (\rho_0 \varphi_z^{-1} / s'_0). \end{aligned} \quad (50)$$

If we use (49) and (50) and linearize, Ertel's Theorem says

$$\begin{aligned} \frac{\partial}{\partial t} \Pi - \nabla \varphi_z^{-1} \cdot \underline{u}_t \times \nabla \varphi - 2\rho_0 s_t \frac{\partial}{\partial z} (\rho_0 \varphi_z^{-1} / s'_0) \\ + 2(s'_0 \rho_0^{-1} \varphi_z)^{-1} \underline{u} \cdot \nabla (s'_0 \rho_0^{-1} \varphi_z) = 0. \end{aligned}$$

Replacing  $\underline{u}_t$  and  $s_t$  in this formula by their values from (5) and (7) one recovers almost at once equation (16).

The function  $\Pi$  is thus approximately the first order part of the "true" potential vorticity  $\rho^{-1} \nabla \times \underline{u} \cdot \nabla s$  - it has been defined somewhat differently in order to obtain a quantity which is constant

in time (in the geostrophically free case) at a fixed place. Although the "true" potential vorticity is no doubt of more basic physical significance, being in fact conserved (following particles) even in the non-linear problem, we have felt it appropriate to give this name to the function  $\bar{\omega}$  which, as we have seen, plays a fundamental role in the linear theory of rotating stratified flow.

This research was partially supported by the Office of Naval Research under Contract Numbers N00014-67-A-0204-0017 and NONR 233(76).

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND

UNIVERSITY OF CALIFORNIA, LOS ANGELES (LNH)

MASSACHUSETTS INSTITUTE OF TECHNOLOGY AND

THE JOHNS HOPKINS UNIVERSITY (WLS)